

## SEQUENCE LISTING

All gene sequences are of 552 base pairs of linear single stranded DNA.

- 5 All protein sequences are of 181 amino acids in length and have enzymatic properties. It should be noted that the sequences presented here are identical to the published lip A sequence except for the first amino acid, methionine present in our sequence. The sequence purified and crystallized in literature does not have the N-terminal methionine.

### 10 GENERAL INFORMATION

APPLICANT: COUNCIL OF SCIENTIFIC AND INDUSTRIAL RESEARCH

TITLE: STABLE GENE VARIANTS OF LIPASES

NUMBER OF SEQUENCES 14

- 15 CORRESPONDING ADDRESS: Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad-500 007, INDIA

Protein Seq ID No. 1

INFORMATION FOR SEQ ID NO:1

- 20 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 181 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: enzyme

- 25 (iii) SEQUENCE DESCRIPTION: SEQ ID NO:1

MAEHNPVVMVHGIGGASFNFAGIKSYLVSQGWSRDKLYAVDFWDKTGTNYNN  
GPVLSRFVQKVLDGTGAKKVDIVAHSMGGANTLYYIKNLDGGNKVANVVTLLG  
ANRLTTGKALPGTDPNQKILYTSIYSSADMIVMNYLSRLDGARNVQIHGGHIGLL  
30 YSSQVNSLIKEGLNGGGQNTN

CORRESPONDING GENE SEQUENCE OF PROTEIN SEQ ID NO:1

- (i) SEQUENCE CHARACTERISTICS:

#pairs (A) LENGTH: 552

- 35 (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

- (iii) HYPOTHETICAL: NO

40

atggctgaacaca atccagtcgt tatggttcacgggtattggag gggcatcatt caattttgcg ggaattaaga gctatctcgt  
atctcagggtggtcgcgagg acaagctgta tgcagttgat ttttgggaca agacaggcac aaattataacaatggaccgg  
tattatcacg atttgtcaa aagggttttag atgaaacggg tgcgaaaaaagtggatattg tcgctcacag catgggggggc  
gcgaacacac ttactacat aaaaaatctggacggcgga ataaagtgc aaacgtcgtg acggttgcg gcgcgaaccg  
45 ttgacgacaggcaaggcgc ttccgggaac agatccaaat caaaagattt tatacacat catttacagc agtgccgata

tgattgtcat gaattactta tcaagattag atggtgctag aaacgttcaaatccatggcg ttggacacat cggccttctg  
tacagcagcc aagtcaacag cctgattaaagaagggtga acggcggggg ccagaatacg aattaatga  
ORGANISM: *Bacillus subtilis*.

IMMEDIATE: NATURAL

5 NAME/KEY: Natural

SEQUENCE ID No.1

Protein Seq ID No. 2

10 INFORMATION FOR SEQ ID NO:1

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 181 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

15 - (ii) MOLECULE TYPE: enzyme

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:2

MAEHNPVVMVHGIGGASFNFAIKSYLSVSGWSRDKLYAVDFWDKTGTNYNN  
20 GPVLSRFVQKVLDDETGAKKVDIVAHSMGGANTLYYIKNLDGGNKVANVVTLLG  
ANRLTTGKALPGTDPNQKILYTSIYSSADMIVMNYLSRLDGARNVQIHGGHIGLL  
YSSQVYSLIKEGLNGGGQNTN

CORRESPONDING GENE SEQUENCE OF PROTEIN SEQ ID NO:2

- (i) SEQUENCE CHARACTERISTICS:

25 #pairs (A) LENGTH: 552

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

30 - (iii) HYPOTHETICAL: NO

atggctgaacaca atccagtcgt tatggttcacggtattggag gggcatcatt caattttgcg ggaattaaga gctatctcgt  
atctcagggctggctgcggg acaagctgta tgcagttgat tttgggaca agacaggcac aaattataacaatggaccgg  
tattatcacg atttgtgcaa aaggttttag atgaacggg tgcgaaaaagtggatattg tcgctcacag catggggggc  
35 gcgaacacac ttactacat aaaaaatctggacggcgga ataaagtgc aaacgtcgtg acggttggcg gcgcgaaccg  
tttgacgacaggcaaggcgc ttccgggaac agatccaaat caaaagattt tatacacatc catttacagc agtgccgata  
tgattgtcat gaattactta tcaagattag atggtgctag aaacgttcaaatccatggcg ttggacacat cggccttctg  
tacagcagcc aagtctacag cctgattaaagaagggtga acggcggggg ccagaatacg aattaatga  
ORGANISM: Artificial

40 IMMEDIATE: Artificial

NAME/KEY: Artificial

SEQUENCE ID No.2

Protein Seq ID No.3

INFORMATION FOR SEQ ID NO:3

- (i) SEQUENCE CHARACTERISTICS:

5 (A) LENGTH: 181 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: enzyme

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:3

10 MAEHNPPVVMVHGIGGASFNFAGIKSYLVSQGWSRDKLYAVDFWDKTGTNYNN  
GPVLSRFVQKVLDGTGVKKVDIVAHSMGGANTLYYIKNLDGGNKVANVVTLGG  
ANRLTTGKALPGTDPNQKILYTSIYSSDDMIVMNYLSRLDGARNVQIHGGHIGLL  
YSSQVYSLIKEGLNGGGQNTN

15 CORRESPONDING GENE SEQUENCE OF PROTEIN SEQ ID NO:3

- (i) SEQUENCE CHARACTERISTICS:

#pairs (A) LENGTH: 552

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

20 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: NO

25 atggctgaacaca atccagtcgt tatggttcacggtattggag gggcatcatt caattttgcg ggaattaaga gctatctcgt  
atctcagggctggtcgcggg acaagctgta tgcagttgat tttgggaca agacaggcac aaattataacaatggaccgg  
tattatcacg atttgtcaa aagggttttag atgaaacggg tgtgaaaaaagtggatattg tcgctcacag catggggggc  
gcgaacacac ttactacat aaaaaatctggacggcgga ataaagtgc aaacgtcgtg acggttggcg gcgcgaaccg  
ttgacgacaggcaaggcgc ttccgggaac agatccaaat caaaagattt tatacacatc catttacagc agtgacgata  
30 tgattgtcat gaattactta tcaagattag atggtgctag aaacgttcaatccatggcg ttggacacat cggccttctg  
tacagcagcc aagtctacag cctgattaagaagggtga acggcggggg ccagaatagc aattaatga  
ORGANISM: Artificial

IMMEDIATE: Artificial

NAME/KEY: Artificial

SEQUENCE ID No.3

35 Protein Seq ID No.4

INFORMATION FOR SEQ ID NO:4

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 181 amino acids

(B) TYPE: amino acid

40 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: enzyme

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:4

MAEHNPPVVMVHGIGGASFNFAGIKSYLVSQGWSRDKLYAVDFWDKTGTNYNN  
GPVLSRFVQKVLDGTGKKVDIVAHSMGGANTLYYIKNLDGGNKVANVVTLGG

ANRLTTGKAPPGTDPNQKILYTSIYSSADMIVMNYLSRLDGARNVQIHGGHIGLL  
YSSQVYSLIKEGLNGGGQNTN

CORRESPONDING GENE SEQUENCE OF PROTEIN SEQ ID NO:4

5 - (i) SEQUENCE CHARACTERISTICS:

- #pairs (A) LENGTH: 552  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10 - (ii) MOLECULE TYPE: DNA

- (iii) HYPOTHETICAL: NO

atggctgaacaca atccagtcgt tatgggtcacgggtattggag gggcattcatt caattttgcg ggaattaaga gctatctcgt  
atctcagggctggctcgcggg acaagctgta tgcagttgat ttitgggaca agacaggcac aaattataacaatggaccgg  
tattatcacg atttgtcaa aaggtttag atgaaacggg taagaaaaagtggatattg tcgctcacag catggggggc  
15 gcgaacacac ttactacat aaaaatctggacggcgga ataaagtgc aaacgtcgtg acggttggcg gcgcgaaccg  
ttgacgacaggcaaggcgct cgggaac agatccaaat caaaagattt tatacacatc cattacagc agtgccgata  
tgattgtcat gaattactta tcaagattag atgggtcgtg aaacgttcaaatccatggcg ttggacacat cggccttctg  
tacagcagcc aagctacag cctgattaaagaagggtga acggcggggg ccagaatacg aattaatga

ORGANISM: Artificial

20 IMMEDIATE: Artificial

NAME/KEY: Artificial

SEQUENCE ID No.4

Protein Seq ID No.5

INFORMATION FOR SEQ ID NO:5

25 - (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 181 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: enzyme

30 (iii) SEQUENCE DESCRIPTION: SEQ ID NO:5

MAEHNPVVMVHGIGGASFNFAIKSYLVSQGWSRDKLYAVDFWDKTGTNYNN  
GPVLSRFVQKVLDETGAKKVDIVAHSMGGANTLYYIKNLDGGNKVANVVTLGG  
ANRLTTGKALPGTDPNQKILYTSIYSSDDMIVMNYLSRLDGARNVQIHGGHIGLL  
YSSQVYSLIKEGLNGGGQNTN

35

CORRESPONDING GENE SEQUENCE OF PROTEIN SEQ ID NO:5

- (i) SEQUENCE CHARACTERISTICS:

- #pairs (A) LENGTH: 552  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

40

- (ii) MOLECULE TYPE: DNA

- (iii) HYPOTHETICAL: NO

atggctgaacaca atccagtcgt tatggttcacgggtattggag gggcatcatt caattttgcg ggaattaaga gctatctcgt  
 atctcagggctggctgcggg acaagctgta tgcagttgat tttggggaca agacaggcac aaattataacaatggaccgg  
 tattatcacg atttgtgcaa aagggttttag atgaaacggg tgcgaaaaaagtggatattg tcgctcacag catggggggc  
 gcgaacacac ttactacat aaaaaatctggacggcgga ataaagtgc aaacgtcgtg acggttggcg gcgcgaaccg  
 5 ttgacgacaggcaaggcgc ttccgggaac agatccaaat caaaagattt tatacacatc catttacagc agtgacgata  
 tgattgtcat gaattactta tcaagattag atggtgctag aaacgttcaaatccatggcg ttggacacat cggccttctg  
 tacagcagcc aagtctacag cctgattaaagaagggtga acggcggggg ccagaatacg aattaatga  
 ORGANISM: Artificial  
 IMMEDIATE: Aritifical  
 10 NAME/KEY: Aritifical  
 SEQUENCE ID No.5

Protein Seq ID No.6  
 INFORMATION FOR SEQ ID NO:6  
 15 - (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 181 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear  
 - (ii) MOLECULE TYPE: enzyme  
 20 (iii) SEQUENCE DESCRIPTION: SEQ ID NO:6  
 MAEHNPPVVMVHIGGGASFNFAGIKSYLVSQGWSRDKLYAVDFWDKTGTNYNN  
 GPVLSRFVQKVLDGTGAKKVDIVAHSMGGANTLYYIKNLDGGNKVANVVTLGG  
 ANRLTTGKAPPGTDPNQKILYTSIYSSDDMIVMNYLSRLDGARNVQIHGGHIGLL  
 YSSQVYSLIKEGLNGGGQNTN  
 25

CORRESPONDING GENE SEQUENCE OF PROTEIN SEQ ID NO:6  
 - (i) SEQUENCE CHARACTERISTICS:  
 #pairs (A) LENGTH: 552  
 (B) TYPE: nucleic acid  
 30 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 - (ii) MOLECULE TYPE: DNA  
 - (iii) HYPOTHETICAL: NO  
 atggctgaacaca atccagtcgt tatggttcacgggtattggag gggcatcatt caattttgcg ggaattaaga gctatctcgt  
 35 atctcagggctggctgcggg acaagctgta tgcagttgat tttggggaca agacaggcac aaattataacaatggaccgg  
 tattatcacg atttgtgcaa aagggttttag atgaaacggg tgcgaaaaaagtggatattg tcgctcacag catggggggc  
 gcgaacacac ttactacat aaaaaatctggacggcgga ataaagtgc aaacgtcgtg acggttggcg gcgcgaaccg  
 ttgacgacaggcaaggcgccct ccgggaac agatccaaat caaaagattt tatacacatc catttacagc agtgacgata  
 40 tgattgtcat gaattactta tcaagattag atggtgctag aaacgttcaaatccatggcg ttggacacat cggccttctg  
 tacagcagcc aagtctacag cctgattaaagaagggtga acggcggggg ccagaatacg aattaatga  
 ORGANISM: Artificial  
 IMMEDIATE: Aritifical  
 NAME/KEY: Aritifical

SEQUENCE ID No.6

Protein Seq ID No.7

INFORMATION FOR SEQ ID NO:7

- (i) SEQUENCE CHARACTERISTICS:

5 (A) LENGTH: 181 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: enzyme

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:7

10

MAEHNPVVMVHIGGGASFNFAGIKSYLSVSGWSDKLYAVDFWDKTGTNYNN  
GPVLSRFVQKVLDDETGAKKVDIVAHSMGGANTLYYIKNLDGGNKVANVVTLGG  
ANRLTTGKALPGTDPNQKILYTSIYSSADMIVMNYLSRLDGASNVQIHGGHIGLL  
YSSQVYSLIKEGLNGGGQNTN

15

CORRESPONDING GENE SEQUENCE OF PROTEIN SEQ ID NO:7

- (i) SEQUENCE CHARACTERISTICS:

#pairs (A) LENGTH: 552

(B) TYPE: nucleic acid

20

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

- (iii) HYPOTHETICAL: NO

25

atggctgaacaca atccagtcgt tatggttcacgggtattggag gggcatcatt caattttgcg ggaattaaga gctatctcgt  
atctcagggcgtggtcgcgagg acaagctgta tgcagttgat tttgggaca agacaggcac aaattataacaatggaccgg  
tattatcacg atttgtcaa aaggtttag atgaaacggg tgcgaaaaaagtggatattg tcgctcacag catggggggc  
gcgaacacac ttactacat aaaaaatctggacggcgga ataaagtgc aaacgtcgtg acggttggcg gcgcgaaccg  
ttgacgacaggcaaggcgc ttccgggaac agatccaaat caaaagattt tatacacatc catttacagc agtgccgata  
tgattgtcat gaattactta tcaagattag atgtgctagtaacgttcaaatccatggcg ttggacacat cggccttctg  
30 tacagcagcc aagtctacag cctgattaaagaagggtga acggcggggg ccagaatacg aattaatga

30

ORGANISM: Artificial

IMMEDIATE: Artificial

NAME/KEY: Artificial

SEQUENCE ID No.7

35

Protein Seq ID No.8

INFORMATION FOR SEQ ID NO:8

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 181 amino acids

40

(B) TYPE: amino acid

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: enzyme

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:8

MAEHNPVVMVHGIGGASFNFAGIKSYLVSQGWSRDKLYAVDFWDKTGTNYNN  
GPVLSRFVQKVLDDETGAKKVDIVAHSMGGANTLYYIKNLDGGNKVANVVTLGG  
ANRLTTGKAPPGTDPNQKILYTSIYSSADMIVMNYLSRLDGARNVQIHGGHIGLL  
YSSQVYSLIKEGLNGGGQNTN

5

CORRESPONDING GENE SEQUENCE OF PROTEIN SEQ ID NO:8

- (i) SEQUENCE CHARACTERISTICS:

#pairs (A) LENGTH: 552

(B) TYPE: nucleic acid

10

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

- (iii) HYPOTHETICAL: NO

15

atggctgaacaca atccagtcgt tatggttcacgggtattggag gggcatcatt caattttgcg ggaattaaga gctatctcgt  
atctcagggtcggtcgcgga acaagctgta tgcagttgat ttttgggaca agacaggcac aaattataacaatggaccgg  
tattatcacg atttgtcaa aagggttttag atgaaacggg tgcgaaaaaagtggatattg tcgctcacag catggggggc  
gcgaacacac ttactacat aaaaaatctggacggcgga ataaagttgc aaacgtcgtg acggttggcg gcgcgaaccg  
ttgacgacaggcaaggcgccct ccgggaac agatccaaat caaaagattt tatacacatc catttacagc agtgccgata  
tgattgtcat gaattactta tcaagattag atggtgctag aaacgttcaaatccatggcg ttggacacat cggccttctg  
20 tacagcagcc aagtctacag cctgattaaagaagggtga acggcggggg ccagaatacg aattaatga  
ORGANISM: Artificial

IMMEDIATE: Artificial

NAME/KEY: Artificial

SEQUENCE ID No.8

25

Protein Seq ID No.9

INFORMATION FOR SEQ ID NO:9

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 181 amino acids

(B) TYPE: amino acid

30

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: enzyme

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:9

35

MAEHNPVVMVHGIGGASFNFAGIKSYLVSQGWSRDKLYAVDFWDKTGTNYNN  
GPVLSRFVQKVLDDETGAKKADIVAHSMGGANTLYYIKNLDGGNKVANVVTLGG  
ANRLTTGKALPGTDPNQKILYTSIYSSADMIVMNYLSRLDGARNVQIHGGHIGLL  
YSSQVYSLIKEGLNGGGQNTN

CORRESPONDING GENE SEQUENCE OF PROTEIN SEQ ID NO:9

- (i) SEQUENCE CHARACTERISTICS:

40

#pairs (A) LENGTH: 552

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

- (iii) HYPOTHETICAL: NO

atggctgaacaca atccagtcgt tatgggtcacgggtattggag gggcatcatt caattttgcg ggaattaaga gctatctcgt  
atctcagggctggctgcggg acaagctgta tgcagttgat tttgggaca agacaggcac aaattataacaatggaccgg  
tattatcacg atttgtcaa aaggttttag atgaaacggg tgcgaaaaaagcgatattg tcgctcacag catggggggc  
5 gcaaacacac ttactacat aaaaaatctggacggcgga ataaagtgc aaacgtcgtg acggttggcg gcgcgaaccg  
ttgacgacaggcaaggcgc ttccgggaac agatccaaat caaaagattt tatacacatc catttacagc agtgccgata  
tgattgtcat gaattactta tcaagattag atggtgctag aaacgttcaaatacatggcg ttggacacat cggccttctg  
tacagcagcc aagtctacag cctgattaaagaagggtga acggcggggg ccagaatacg aattaatga

ORGANISM: Artificial

10 IMMEDIATE: Artificial

NAME/KEY: Artificial

SEQUENCE ID No.9

Protein Seq ID No.10

15 INFORMATION FOR SEQ ID NO:10

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 181 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

20 - (ii) MOLECULE TYPE: enzyme

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:10

MAEHNPPVVMVHGIGGASFNFAIKSYLVSQGWSDKLYAVDFWDKTGTNYNN  
GPVLSRFVQKVLDETGAKKVDIVAHSMGGANTLYYIKNLDGGNKVANVVTLLG  
ANRLTTGKALPGTDPNQKILYTSIYSSADMIVMNYLSRLVGARNVQIHGGHIGLL  
25 YSSQVYSLIKEGLNGGGQNTN

CORRESPONDING GENE SEQUENCE OF PROTEIN SEQ ID NO:10

- (i) SEQUENCE CHARACTERISTICS:

#pairs (A) LENGTH: 552

30 (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

- (iii) HYPOTHETICAL: NO

35 atggctgaacaca atccagtcgt tatgggtcacgggtattggag gggcatcatt caattttgcg ggaattaaga gctatctcgt  
atctcagggctggctgcggg acaagctgta tgcagttgat tttgggaca agacaggcac aaattataacaatggaccgg  
tattatcacg atttgtcaa aaggttttag atgaaacggg tgcgaaaaaagtgatattg tcgctcacag catggggggc  
gcaaacacac ttactacat aaaaaatctggacggcgga ataaagtgc aaacgtcgtg acggttggcg gcgcgaaccg  
ttgacgacaggcaaggcgc ttccgggaac agatccaaat caaaagattt tatacacatc catttacagc agtgccgata  
40 tgattgtcat gaattactta tcaagatta gttggtgctag aaacgttcaaatacatggcg ttggacacat cggccttctg  
tacagcagcc aagtctacag cctgattaaagaagggtga acggcggggg ccagaatacg aattaatga

ORGANISM: Artificial

IMMEDIATE: Artificial



NAME/KEY: Aritifical

SEQUENCE ID No.10

Protein Seq ID No.11

INFORMATION FOR SEQ ID NO:11

5 - (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 181 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: enzyme

10 (iii) SEQUENCE DESCRIPTION: SEQ ID NO:11

MAEHNPPVVMVHGIGGASFNFAGIKSYLVSQGWSRDKLYAVDFWDKTGTNYNN  
GPVLSRFVQKVLDDETGVKKVDIVAHSMGGANTLYYIKNLDGGNKVANVVTLGG  
ANRLTTGKALPGTDPDQKILYTSIYSSDDMIVMNYLSRLDGARNVQIHGGHIGLL  
15 YSSQVYSLIKEGLNGGGQNTN

CORRESPONDING GENE SEQUENCE OF PROTEIN SEQ ID NO:11

- (i) SEQUENCE CHARACTERISTICS:

#pairs (A) LENGTH: 552

20 (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

- (iii) HYPOTHETICAL: NO

25 atggctgaacaca atccagtcgt tatgggtcacggtattggag gggcatcatt caattttgcg ggaattaaga gctatctcgt  
atctcagggctggcgcgagg acaagctgta tgcagttgat tttgggaca agacaggcac aaattataacaatggaccgg  
tattatcacg atttgtcaa aaggttttag atgaaacggg tgtgaaaaaagtggatattg tcgctcacag catggggggc  
gcgaacacac ttactacat aaaaaatctggacggcgga ataaagtgc aaacgtcgtg acggttgcg gcgcgaaccg  
30 ttgacgacaggcaaggcgc ttccgggaac agatccagat caaaagattt tatacacatc catttacagc agtgaegata  
tgattgtcat gaattactta tcaagattag atgggtctag aaacgttcaaatccatggcg ttggacacat cggccttctg  
tacagcagcc aagtctacag cctgattaaagaagggctga acggcggggg ccagaatacg aattaatga

ORGANISM: Artificial

IMMEDIATE: Aritifical

NAME/KEY: Aritifical

35 SEQUENCE ID No.11

Protein Seq ID No.12

INFORMATION FOR SEQ ID NO:12

- (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 181 amino acids

40 (B) TYPE: amino acid

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: enzyme

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:12

MAEHNPVVMVHIGGASFNFAGIKSYLVSQGWSRDKLYAVDFWDKTGTNYNN  
GPVLSRFVQKVLDGTGAKKVDIVAHSMGGANTLYYIKNLDGGNKVANVVTLGG  
ANRLTTGKAPPGTDPDQKILYTSIYSSDDMIVMNYLSRLDGARNVQIHGGHIGLL  
YSSQVYSLIKEGLNGGGQNTN

5

CORRESPONDING GENE SEQUENCE OF PROTEIN SEQ ID NO:12

- (i) SEQUENCE CHARACTERISTICS:

#pairs (A) LENGTH: 552

(B) TYPE: nucleic acid

10 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

- (iii) HYPOTHETICAL: NO

15 atggctgaacaca atccagtcgt tatggttcacgggtattggag gggcatcatt caattttgcg ggaattaaga gctatctcgt  
atctcagggtggtcgcgagg acaagctgta tgcagttgat tttgggaca agacaggcac aaattataacaatggaccgg  
tattatcacg atttgtcaa aaggttttag atgaaacggg tgcgaaaaaagtggatattg tcgctcacag catggggggc  
gcgaacacac ttactacat aaaaaatctggacggcgga ataaagtgc aaacgtcgtg acggttggcg gcgcgaaccg  
ttgacgacaggcaaggcgccet ccgggaac agatccagat caaaagattt tatacacatc catttacagc agtgacgata  
20 tgattgtcat gaattactta tcaagattag atggtgctag aaacgttcaaattccatggcg ttggacacat cggccttctg  
tacagcagcc aagtctacag cctgattaagaagggtga acggcggggg ccagaatacg aattaatga

ORGANISM: Artificial

IMMEDIATE: Artificial

NAME/KEY: Artificial

SEQUENCE ID No.12

25 INFORMATION FOR SEQ ID NO:13

- (i) SEQUENCE CHARACTERISTICS:

#pairs (A) LENGTH: 24

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

30 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA

(iii) HYPOTHETICAL: NO

5'-CGCCAGGGTTTTCCCAGTCACGAC-3'

35 ORGANISM: Artificial

IMMEDIATE: Artificial

NAME/KEY: Oligonucleotide

SEQUENCE ID No.13

INFORMATION FOR SEQ ID NO:14

40 - (i) SEQUENCE CHARACTERISTICS:

#pairs (A) LENGTH: 22

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
- (ii) MOLECULE TYPE: DNA  
(iii) HYPOTHETICAL: NO

5

5'-TGACACAGGAAACAGCTATGAC-3'

ORGANISM: Artificial

IMMEDIATE: Artificial

NAME/KEY: Oligonucleotide

10

SEQUENCE ID No.14